BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 10: e183020
doi: 10.3897/biss.10.183020

OPEN ACCESS

Conference Abstract

Adapting a Globally-Trained Plant Identification Model for Multi-Species Detection of Invasive Alien Plants in Roadside Imagery

Hervé Goëau[‡], Vincent Espitalier[§], Jean-Christophe Lombardo[|], Christophe Botella[|], Toke Thomas Høye[¶], Mads Dyrmann[#], Pierre Bonnet[§], Alexis Joly[|]

‡ Cirad, UMR AMAP, Montpellier, France
§ CIRAD, UMR AMAP, Montpellier, France
| Inria, LIRMM, Univ Montpellier, CNRS, Montpellier, France
¶ Aarhus University, Aarhus, Denmark
# AI Lab ApS, Aarhus, Denmark

Corresponding author: Hervé Goëau (herve.goeau@cirad.fr)

Received: 19 Dec 2025 | Published: 08 Jan 2026

Citation: Goëau H, Espitalier V, Lombardo J-C, Botella C, Høye TT, Dyrmann M, Bonnet P, Joly A (2026) Adapting a Globally-Trained Plant Identification Model for Multi-Species Detection of Invasive Alien Plants in Roadside Imagery. Biodiversity Information Science and Standards 10: e183020.
https://doi.org/10.3897/biss.10.183020

Abstract

Invasive alien plants drive biodiversity loss and ecosystem disruption, costing billions globally. Roadsides faciliate spread, yet large-scale monitoring is difficult. Traditional methods like object detection and instance segmentation require costly, time-consuming annotations, impractical for thousands of species. Citizen science platforms like Pl@ntNet use efficient classification models for global identification, yet they struggle with high-resolution multi-species imagery. In this work we adapted the Pl@ntNet global model, trained on millions of images, to detect multiple invasive species in roadside imagery (Espitalier et al. 2025).

Pl@ntNet uses a Vision Transformer trained on 6.6 million images covering 43,683 species. Optimized for single-species classification from fixed-resolution (384×384) close-ups of plant organs, it fails on high-resolution complex, multi-scale roadside scenes where plant size varies.

© Goëau H et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

We explored two strategies to bridge this domain gap. VaMIS (Variable Model Input Size), forces the model to process the whole roadside scene at once at a larger but non optimal resolution (1024×768). Multi-scale tiling acts like a scanner, decomposing the high-resolution image into hundreds of overlapping patches at multiple zoom levels. Each patch is analyzed independently at the model's native optimal resolution (384×384). If a species is identified with high confidence in any patch, it is considered present in the full image.

Both were tested on the Danish Road Dataset (Dyrmann et al. 2021), containing 14,808 high-resolution (4024×3036) vehicle-mounted images for evaluating the detection of six invasive taxa: *Cytisus scoparius*, *Lupinus polyphyllus*, *Pastinaca sativa*, *Reynoutria* spp., *Rosa rugosa*, and *Solidago* spp. 44% of images contain no invasive species, and few include multiple taxa. Performance was measured using balanced accuracy (average recall of presence/absence classes, mitigating imbalance) and AUC (Area Under the ROC Curve, assessing global discrimination of True vs. False Positive rates). F1 and Jaccard index were used on a multi-label test set of 206 images.

The tiling approach demonstrated strong generalization, achieving 84.9% balanced accuracy and 91.6% AUC, despite requiring no additional annotations, whereas VaMIS underperformed. To reduce domain shift and re-optimize internal parameters, essential for VaMIS to handle full scenes, we fine-tuned the whole transformer (VaMIS) or the classification head (tiling). Fine-tuning improved both: VaMIS rose from 66.8 to 91.8% accuracy, and tiling from 84.9 to 92.1%. On the multi-species test set, fine-tuned tiling achieved the best results (F1 = 0.85, Jaccard = 0.76), outperforming fine-tuned VaMIS and baseline CNN or YOLO, both models representing former classification and object detection baselines reproduced from previous work (Dyrmann et al. 2021) for comparison.

Performance varied by species, but detection of *Rosa rugosa* consistently exceeded 99% AUC even without fine-tuning. Tiling was about 34x slower but far less resource-intensive to fine-tune (20 min vs 25 h).

These results highlight complementary strategies. Tiling offers a generic, training-free pipeline, while VaMIS is faster but needs fine-tuning and memory. Reusing Pl@ntNet models cuts annotation costs, enabling scalable monitoring. Together, these strategies extend biodiversity monitoring to realistic roadside imagery, aligning with FAIR principles.

Keywords

invasive species, roadside biodiversity monitoring, plant recognition, deep learning, large-scale species identification, tiling, Pl@ntNet

Presenting author

Hervé Goëau

Presented at

Living Data 2025

Acknowledgements

This work was supported by the European Commission through the GUARDEN (101060693) and MAMBO (101060639) projects. This work was provided with computer and storage resources by GENCI at IDRIS thanks to the grant 2024-A0171011389 on the supercomputer Jean Zay's V100, A100 and H100 partitions.

Funding program

This work was supported by the European Commission through the GUARDEN (101060693) and MAMBO (101060639) projects.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Dyrmann M, Mortensen AK, Linneberg L, Høye TT, Bjerge K (2021) Camera Assisted Roadside Monitoring for Invasive Alien Plant Species Using Deep Learning. Sensors (Basel, Switzerland) 21 (18). https://doi.org/10.3390/s21186126
- Espitalier V, Lombardo J, Goëau H, Botella C, Høye TT, Dyrmann M, Bonnet P, Joly A (2025) Adapting a global plant identification model to detect invasive alien plant species in high-resolution road side images. Ecological Informatics 89 https://doi.org/10.1016/j.ecoinf.2025.103129